The Relationship between sustainability performance and sustainability disclosure – Reconciling voluntary disclosure theory and legitimacy theory

Hummel, Katrin ; Schlick, Christian

Abstract: The relationship between sustainability performance and sustainability disclosure remains ambiguous, both theoretically and empirically. Voluntary disclosure theory would suggest that the relationship should be positive, whereas legitimacy theory points toward a negative relationship. However, the empirical evidence regarding this relationship is mixed, which indicates that the two theories are not necessarily contradictory but that they are instead two sides of the same coin. This paper refines the theoretical reasoning associated with the two theories and provides empirical evidence for their reconciliation by moving the focus of inquiry from the quantity of sustainability disclosure toward its quality. Our results reveal that – consistent with voluntary disclosure theory – superior sustainability performers choose high-quality sustainability disclosure to signal their superior performance to the market. In addition, based on legitimacy theory, poor sustainability performers prefer low-quality sustainability disclosure to disguise their true performance and to simultaneously protect their legitimacy. The results remain robust to various additional analyses. Thus, the paper indicates that the two theories dovetail with one another by redirecting the focus toward the quality of sustainability disclosure.

DOI: https://doi.org/10.1016/j.jaccpubpol.2016.06.001

The relationship between sustainability performance and sustainability disclosure – Reconciling voluntary disclosure theory and legitimacy theory

Katrin Hummel, Christian Schlick

Journal of Accounting and Public Policy, forthcoming

The article is available at:

http://www.sciencedirect.com/science/article/pii/S0278425416300333

ABSTRACT

The relationship between sustainability performance and sustainability disclosure remains ambiguous, both theoretically and empirically. Voluntary disclosure theory would suggest that the relationship should be positive, whereas legitimacy theory points toward a negative relationship. However, the empirical evidence regarding this relationship is mixed, which indicates that the two theories are not necessarily contradictory but that they are instead two sides of the same coin. This paper refines the theoretical reasoning associated with the two theories and provides empirical evidence for their reconciliation by moving the focus of inquiry from the quantity of sustainability disclosure toward its quality. Our results reveal that – consistent with voluntary disclosure theory – superior sustainability performers choose high-quality sustainability disclosure to signal their superior performance to the market. In addition, based on legitimacy theory, poor sustainability performers prefer low-quality sustainability disclosure to disguise their true performance and to simultaneously protect their legitimacy. The results remain robust to various additional analyses. Thus, the paper indicates that the two theories dovetail with one another by redirecting the focus toward the quality of sustainability disclosure.

Keywords: sustainability disclosure, sustainability performance, legitimacy theory, voluntary disclosure theory

ACKNOWLEDGEMENTS

We would like to express our thanks to Editor-in-Chief Lawrence A. Gordon, Editor Martin Loeb, two anonymous reviewers, Dieter Pfaff, Hui Chen, and Patricia Ruffing as well as participants at several University of Zurich PhD seminars, the 3rd ERMAC conference in Vienna, the 37th European Accounting Association Annual Congress in Tallinn and the 76th VHB Annual Congress in Leipzig for their helpful comments and feedback on previous versions of this paper. Moreover, we thank RepRisk for providing access to its ESG Risk Platform. All remaining errors are our own.

1. Introduction

Previous research has not yet established a consistent understanding regarding the relationship between sustainability performance and sustainability disclosure. In essence, two theoretical concepts are involved. On the one hand, voluntary disclosure theory predicts that a company with good sustainability performance is incentivized to disclose information regarding its performance to increase its market value. This stream of research posits a positive relationship between sustainability performance and the quantity of sustainability disclosure (i.e., superior sustainability performers disclose more). On the other hand, legitimacy theory argues that companies employ sustainability disclosure to improve the public perception of their sustainability performance (Deegan, 2002). Researchers therefore interpret a negative relationship between sustainability performance and the quantity of sustainability disclosure (i.e., poor sustainability performers disclose more) as an indication of the applicability of legitimacy theory (Cho et al., 2012; Patten, 2002). Thus, these two theories yield opposing predictions regarding the relationship between sustainability performance and sustainability disclosure, and the mixed empirical results from prior studies have not yet clarified this relationship (for a positive relationship between sustainability performance and sustainability disclosure, see Al-Tuwaijri et al., 2004; Clarkson et al., 2008; for a negative relationship, see Cho and Patten, 2007; de Villiers and van Staden, 2006).

Recent research has therefore inquired whether these two theories are not mutually exclusive but are instead two sides of the same coin and has found some preliminary evidence to justify this line of analysis. For instance, Clarkson et al. (2008) ascribe a positive relationship between environmental performance and environmental disclosure as evidence for the application of voluntary disclosure theory but refer to legitimacy theory to explain “interesting patterns in the data” (Clarkson et al., 2008). They call for a switch in the "focus of enquiry" of future environmental disclosure research to investigate the concurrent applicability of the two theories more rigorously (Clarkson et al., 2008).

Whereas previous studies focus primarily on the *quantity* of sustainability disclosure by classifying disclosure items as either disclosed or non-disclosed, future research must illuminate *how* information is disclosed. In addition, to precisely assess the theoretical implications derived from voluntary disclosure theory and legitimacy theory, proxies for sustainability disclosure and sustainability performance must capture similar content. However, prior research

is characterized by a variety of different approaches to measuring sustainability performance that range from the use of single indicators of environmental performance, such as emissions or waste (Clarkson et al., 2011), to rating metrics provided by specialized rating agencies, such as Kinder, Lydenberg, Domini (KLD, today MSCI) (Cho and Patten, 2007; Cho et al., 2006; Dawkins and Fraas, 2011).

Taken together, the measurement of both sustainability disclosure and sustainability performance is essential to reconcile the two theories. This paper therefore provides refined measurement approaches for sustainability disclosure and sustainability performance. With respect to measuring sustainability disclosure, we focus on the *quality* – rather than the *quantity* – of sustainability disclosure. While the (mandatory) financial disclosure literature, in particular, is concerned with the quality of reported *earnings* (for a literature review, see Beyer et al., 2010; Leuz and Wysocki, 2008), any metric used to measure the quality of sustainability disclosure must account for its voluntary nature and cover a broader spectrum of information. We therefore concentrate on the reporting quality of 14 disclosure items in the environmental and social dimensions of sustainability. In contrast to earnings quality studies, our measure for high-quality disclosure does not relate to the ex post truthfulness of the disclosed information but instead accounts for traditional disclosure quality criteria such as verifiability, reliability, comparability and consistency (Leuz and Wysocki, 2008, p. 25). We argue that only high-quality reporting of quantitative sustainability information allows outsiders to assess the true sustainability performance of a company. Our measurement of sustainability performance is based on manually collected data regarding four environmental and four social performance indicators to ensure content-based congruence between the measurement of sustainability disclosure and sustainability performance. The data are rescaled on an industry-group basis and aggregated into an overall sustainability performance score.

We posit two hypotheses to test the applicability of voluntary disclosure theory and legitimacy theory separately. On the one hand, we expect to find a positive relationship between a firm's sustainability performance and high-quality sustainability disclosure. This hypothesis reflects the underlying reasoning of voluntary disclosure theory that a company with superior sustainability performance voluntarily discloses sustainability information to increase its market value (Clarkson et al., 2008). We argue that this reasoning applies primarily to high-quality sustainability disclosure because only high-quality disclosure allows outside investors

to assess a company's true sustainability performance. On the other hand, we expect to find a negative relationship between a firm's sustainability performance and low-quality sustainability disclosure. Legitimacy theory suggests that particularly poorly performing companies use sustainability disclosure as a legitimation strategy to influence public perceptions of their sustainability performance (Deegan, 2002; O'Donovan, 2002; Sethi, 1978). We argue that these companies prefer to disclose low-quality information – information that is opaque, incomplete or superficial – to obscure their poor sustainability performance while simultaneously attempting to maintain legitimacy.

The results from a regression analysis for a sample of 195 European companies support our notion that rather than being competitive and mutually exclusive, the two theories instead simultaneously explain the reporting quality of sustainability information. We provide evidence that superior sustainability performers choose high-quality sustainability reporting to signal their superior performance to the market. On the other hand, poor sustainability performers provide low-quality sustainability information to disguise their true performance while simultaneously attempting to maintain their legitimacy. The results from several model variations and supplemental analyses support the robustness of our findings.

Our study makes several contributions to the literature. First, to the best of our knowledge, this is the first study to investigate the applicability of *both* voluntary disclosure theory *and* legitimacy theory in explaining the relationship between sustainability performance and sustainability disclosure. Contrary to prior research, we do not assume that these theories are mutually exclusive but argue instead that the two theories can be reconciled. By redirecting the focus of enquiry from sustainability disclosure quantity to sustainability disclosure quality, we present a research setting in which we can empirically assess our hypotheses. Second, using a sample of 195 European companies, we provide robust empirical evidence that supports our reasoning. In addition, our results augment our knowledge about other determinants of the quality of sustainability disclosure, which has not been sufficiently understood in Western Europe (Fifka, 2013). Third, we develop improved and transparent approaches to measuring sustainability performance and the quality of sustainability disclosure that include both environmental and social dimensions. The application of these measurement schemes might be useful to future research in this field. From a practical perspective, our findings highlight

the need for precise and binding disclosure standards for core quantitative sustainability information in Europe.

The remainder of this paper is structured as follows. Section 2 reviews the related literature and develops the hypotheses. Section 3 explains our research design and focuses, in particular, on the measurement schemes for sustainability performance and the quality of sustainability disclosure. In addition, this section describes the data sample and the empirical model. Section 4 provides descriptive results and our findings from regression analyses and robustness checks. Section 5 concludes the paper.

2. Literature review and hypotheses development

From a theoretical perspective, nearly all of the previous empirical studies on the relationship between sustainability performance and sustainability disclosure are based on either voluntary disclosure theory or legitimacy theory (Al-Tuwaijri et al., 2004; Cho et al., 2012; Cho and Patten, 2007; Clarkson et al., 2008; Clarkson et al., 2011; de Villiers and van Staden, 2006; Patten, 2002). With respect to voluntary disclosure theory, the "unraveling" of private information serves as the baseline model for corporate voluntary disclosure. This unraveling result is subject to a number of conditions, which include, in particular, costless and truthful disclosure.

However, because disclosure is *not* costless, rational managers withhold unfavorable information below a critical threshold disclosure level (Verrecchia, 1983). Although this theory originally referred exclusively to the voluntary disclosure of *financial* information, researchers have also applied it to explain the voluntary disclosure of *non-financial* information (Bewley and Li, 2000; Clarkson et al., 2008; Li et al., 1997) by arguing that a company with superior sustainability performance voluntarily discloses non-financial information to reveal the nature of its true performance and to (potentially) increase its market value (Clarkson et al., 2008). Such value-increasing effects of sustainability disclosure are documented in the literature in different settings (Clarkson et al., 2013; De Villiers and Marques, 2016; Dhaliwal et al., 2011; Plumlee et al., 2015). For instance, Dhaliwal et al. (2011) show that superior sustainability performers have significantly lower costs of equity capital when they publish a standalone sustainability report for the first time. Moreover, such initiating firms with superior sustainability performance attract more dedicated institutional investors and analyst cover-

age. Similarly, in a multi-country setting, De Villiers and Marques (2016) reveal that the quantity of sustainability disclosure is positively correlated with higher stock prices.

Another important condition of the unraveling result is the truthfulness of disclosure. With respect to voluntary financial disclosure, this assumption is typically justified by the litigation and reputational risks associated with untruthful reporting (Verrecchia, 2001). Similar reasoning applies to the voluntary disclosure of sustainability information, particularly with regard to public reputation and the relevance of a sustainable image for corporate success (Ameer and Othman, 2012; Wood, 1991). In addition, the growing number of externally assured sustainability reports limits the possibility of misrepresentation (KPMG, 2011). Under the assumption that untruthful disclosure is unlikely and in light of a lack of precise and binding sustainability reporting standards, companies have substantial leeway in determining both the quantity and the quality of sustainability disclosure. Previous research (Al-Tuwaijri et al., 2004; Bewley and Li, 2000; Cho and Patten, 2007; Clarkson et al., 2008; Clarkson et al., 2011) provides valuable insights into the quantity of information provided. However, firms may provide high-quality information regarding topics that are favorable to themselves (superior performance) while disclosing only low-quality information on topics that may be detrimental to their interests (poor performance). To overcome this potential bias, we concentrate specifically on the disclosure *quality* of core sustainability information.

Our definition of high-quality disclosure draws on "desirable properties of […] financial reports", such as verifiability, reliability, comparability, and consistency (Leuz and Wysocki, 2008). We define high-quality disclosure as the complete disclosure of relevant and comparable numerical data that fulfill or exceed clearly defined quality requirements. Low-quality disclosure refers to any other information that does not fulfill the criteria for high-quality disclosure, irrespective of its quantity. We argue that the reasoning of voluntary disclosure theory applies primarily to high-quality disclosure because such disclosure offers the transparency necessary to be both reliable and comparable to disclosure by other firms. Companies with superior sustainability performance regarding a specific performance indicator prefer to disclose high-quality information to signal their true (unobservable) performance type. Moreover, such disclosure cannot be easily mimicked by companies with poor sustainability performance (Clarkson et al., 2008), which enables superior performers to distinguish themselves from poor performers. Following this reasoning, we formally state our hypothesis as follows:

H1: There is a positive relationship between corporate sustainability performance and high-quality corporate sustainability disclosure.

Legitimacy theory offers another theoretical explanation for the voluntary disclosure of non-financial information. Suchman (1995) defines legitimacy as "a generalized perception or assumption that the actions of an entity are desirable, proper, or appropriate within some socially constructed system of norms, values, beliefs, and definitions". This abstract concept of society is more precisely delineated by Freeman's (1984) definition of stakeholders as "groups […] that can affect, or are affected by, the accomplishment of organizational purpose" (Wood, 1991).

If a firm's legitimacy is threatened because stakeholders perceive its performance as non-sustainable, the long-term survival of the firm is at risk (Davis, 1973). Such negative effects may originate from poor image, customer dissatisfaction, hiring issues, litigation and stricter regulation, among other causes (Ameer and Othman, 2012; Wood, 1991). Legitimacy theory suggests that particularly poorly performing companies use sustainability disclosure as a legitimation tactic to influence public perceptions regarding their sustainability performance (Deegan, 2002; O'Donovan, 2002; Sethi, 1978). Thus, companies that perform poorly on a specific performance indicator prefer low-quality information – information that is superficial, incomplete, not easily subject to comparison or ambiguous – to obscure their poor true state while simultaneously maintaining legitimacy creating a proper sustainability image. Consequently, we posit the following relationship:

H2: There is a negative relationship between corporate sustainability performance and low-quality corporate sustainability disclosure.

Our reasoning that posits that the two theories are not mutually exclusive is reflected in the two hypotheses that address two aspects (i.e., high-quality and low-quality) of the same sustainability disclosure. In our quality setting, firms' reporting behavior is assumed to be driven by incentives aimed at increasing market value (voluntary disclosure theory) *and at the same time* by attempts to avoid the negative consequences of threatened legitimacy (legitimacy theory). By testing the hypotheses simultaneously and not against one another, the research setting allows a distinction between the applicability of each theory. Consistent with both voluntary disclosure theory and legitimacy theory, companies may also opt for non-disclosure re-

garding a specific issue. Thus, evidence falsifying one of the hypotheses does not necessarily allow inferences for the other hypothesis.

3. Research design

3.1. Measurement of corporate sustainability disclosure quality

According to Leuz and Wysocki (2008), there is neither a common understanding of nor a common measurement approach to "high-quality" financial disclosure. Typical measures for the quality of *mandatory* financial disclosure are based on the properties of reported earnings, such as earnings smoothing, earnings persistence or earnings value-relevance. The *voluntary* financial disclosure literature is characterized by a variety of different measurement approaches, such as rankings and self-constructed content-based measures. Such content-based measures are also the primary research methodology in analyses of voluntary *non-financial* disclosure. Researchers first identify relevant information items and then assess the respective disclosures for each item (Al-Tuwaijri et al., 2004; Cho et al., 2012; Cho and Patten, 2007; Cho et al., 2006; Clarkson et al., 2008; Clarkson et al., 2011; de Villiers and van Staden, 2006; Hughes et al., 2001; Patten, 2002; Wiseman, 1982).

Many of these studies distinguish between “monetary” or “hard” and “non-monetary” or “soft” disclosure items (Cho et al., 2012; Cho and Patten, 2007; Clarkson et al., 2008; Clarkson et al., 2011; Patten, 2002). Hard disclosure items focus on a firm’s disclosures related to environmental performance indicators such as data on emissions, water use and recycling, whereas soft disclosure items concentrate on the disclosure of a firm’s vision, environmental strategy and commitment to environmentally responsible management. The disclosure items are typically observed on a binary basis (disclosure vs. non-disclosure) and aggregated into an overall disclosure score. In so doing, the aggregated scores mostly refer to the *quantity* of disclosure but are not intended to capture the *quality* of disclosure, i.e., the shades of reporting between disclosure and non-disclosure.

Other measurement schemes are based on ordinal ratings and thus attempt to capture disclosure quality directly. In this case, a higher rating is typically assigned to quantitative disclosure, and a lower rating is assigned to non-quantitative disclosure (Aerts and Cormier, 2009; Aerts et al., 2008; Al-Tuwaijri et al., 2004; Wiseman, 1982). Because the overall disclosure

score consists of a mixture of quantitative and qualitative ratings for each item, both voluntary disclosure theory and legitimacy theory predict a positive relationship between sustainability performance and sustainability disclosure (Clarkson et al., 2011).[1] Again, it is difficult to clearly distinguish between the applicability of each theory.

Against this background, we provide a measurement scheme for sustainability disclosure that concentrates on its *quality* as opposed to its *quantity*. Our measurement scheme for sustainability disclosure quality draws on the "desirable properties" of high-quality financial disclosure: verifiability, reliability, comparability, and consistency (Leuz and Wysocki, 2008). These disclosure properties are translated into our measurement scheme for sustainability disclosure quality that integrates the environmental as well as the social dimension of sustainability, in particular, employee-related information.[2] Our index of disclosure items is closely linked to the reporting requirements defined by the GRI sustainability reporting guidelines version 3.1, which are considered the most commonly used international sustainability reporting standards at present (Ballou et al., 2006; Gray, 2006; KPMG, 2011). For each disclosure item, the GRI guidelines provide precise descriptions of all material information that must be disclosed by a company. Each sustainability category – environmental and social – consists of seven performance indicators. All indicators in our scheme are classified by the GRI guidelinees as core indicators that are generally applicable to most companies (GRI, 2011a). Thus, in contrast to the measurement of sustainability disclosure *quantity*, our measurement of sustainability disclosure *quality* is not comprehensive but instead concentrates on core sustainability disclosure items. Because these core performance indicators relate to fundamental aspects of a company's sustainability performance, we argue that the disclosure quality of these indicators is an adequate proxy for the (overall) quality of a company's sustainability disclosure.

---

[1] According to voluntary disclosure theory, better environmental performance is related to more quantitative disclosure (higher disclosure scores), according to legitimacy theory worse environmental performance is related to less quantitative and more non-quantitative disclosure (lower disclosures scores).

[2] We do not include the economic dimension of sustainability because the reporting of key economic information is mandatory and subject to both national and international accounting standards. Moreover, the investigation of the various aspects of mandatory and voluntary financial disclosure represents its own line of research (for an overview, see Beyer et al., 2010; Leuz and Wysocki, 2008).

Insert Table 1 about here

Table 1 provides an overview of the measurement scheme. For each of the respective disclosure items, exactly one point is awarded for high-quality disclosure, low-quality disclosure or non-disclosure. We define high-quality disclosure as the disclosure of numerical data on a company-wide level that fulfill or exceed the minimum requirements derived from the GRI guidelines G3.1 and described in Table 1. If these requirements are not fulfilled and any other information regarding the respective indicator is provided, one point is awarded for low-quality disclosure. Even if a firm provides extensive information on an indicator but withholds the data that are truly relevant according to the G3.1 guidelines, the disclosure is coded as low-quality. This approach ensures that our measurement scheme is not biased by obfuscating disclosure strategies as implied by legitimacy theory. When there is no information at all, the item is classified as not reported. We include sector-specific adjustments for high-quality requirements in the environmental category to account for industry-specific variations in the relevance of specific disclosure items. High-quality (low-quality) disclosure scores are calculated as the sum of all high-quality (low-quality) disclosure item scores and range between 0 (minimum) and 14 (maximum).

Because the indicator-specific differentiation between high-quality disclosure, low-quality disclosure and non-disclosure refers to the same disclosure items, our measurement scheme for sustainability disclosure quality enables an empirically testable distinction between the applicability of voluntary disclosure theory (high-quality information) and legitimacy theory (low-quality information). An earlier version of the measurement scheme was pre-tested by both authors and then adjusted for additional details to ensure homogeneity in coding results across different coders. The clear and detailed descriptions of the disclosure items support the reliability of our measure. According to the triple-bottom-line approach developed by Elkington (1997), corporate sustainability is a multi-dimensional construct. Therefore, including multiple disclosure items from both the environmental and social dimensions of sustainability strengthens our measurement scheme's validity. As with all content-based measures of disclosure quality, our measure of the quality of sustainability disclosure relies on the assumption of truthful disclosure. In section 4.3, we therefore provide additional analyses to check the robustness of our results with respect to this assumption.

3.2. Measurement of corporate sustainability performance

Prior research on the relationship between sustainability performance and sustainability disclosure has employed a variety of different measurement approaches for sustainability performance. Some studies are based on one or two indicators – such as emissions or waste – as proxies for overall environmental performance (Al-Tuwaijri et al., 2004; Clarkson et al., 2008). These measurement approaches involve important aspects of sustainability performance and have paved the way for more complex proxies that cover several performance indicators to increase validity (Horváthová, 2012). Other studies utilize rating metrics provided by external agencies (Cho and Patten, 2007; Cho et al., 2006; Dawkins and Fraas, 2011). One of the most widely used ratings is the KLD (today, the MSCI) rating, which is based on binary data covering 14 items, through which the environmental performance of a company is assessed in terms of strengths and weaknesses (Chatterji et al., 2009).

Both the coverage of different dimensions of sustainability and the large dataset of companies covered by the rating makes it attractive for researchers. However, neither the criteria nor the threshold levels for these binary assessments are revealed, and the ratings process is thus not fully transparent. Therefore, the reliability cannot be evaluated from an outsider's perspective. Moreover, the binary assessment insufficiently accounts for substantial variances in the underlying data, which may also influence the validity of the proxy. As a consequence, researchers have repeatedly questioned the appropriateness of the KLD database for purposes of academic research (Chin et al., 2013; Chiu and Sharfman, 2011) and have called for the development of an improved measure of sustainability performance (Hong and Andersen, 2011).

We heed this call and develop a more refined measure of sustainability performance. Our measurement scheme consists of four environmental and four social performance indicators. We directly refer to the data provided by the respective company. If data are provided only for certain countries, business areas or employee groups but cover at least 80% of total sales or total employees, we extrapolate the data to the entire corporation. If data are unreported, the respective performance indicator for the company contains a missing value, and companies with more than two missing values in one dimension are excluded from the sample. Note that

performance indicators may contain valid values even when the corresponding disclosure items are of low quality.[3] Table 2 provides an overview of the performance indicators, the measurements of these indicators, the units of measurement and the ing sustainability disclosure items.

Insert Table 2 about here

The original data for each indicator are arranged by industry groups and then winsorized within each industry group at the top and bottom tails at a 10% level to limit the influence of outliers (Tukey, 1962). Next, all values are transformed into a continuous [0, 1] scale per industry group by assigning "0" to the worst and "1" to the best performance indicator values and by rescaling all other values proportionally. Thus, this step allows all performance indicators with their differing measurement units to be aggregated into a total performance score. A minimum of five companies per industry group is required to define reasonable peer groups. Rescaling on an industry group basis allows for the comparability of performance scores across different industry groups. Our final sustainability performance score is calculated as the arithmetic mean of the means of the environmental and social performance indicators. Each performance indicator is therefore weighted equally, and missing values are approximated by the average of the remaining indicators within each dimension. Potential biases resulting from this approach are analyzed in the "Supplemental analyses" section of this paper (section 4.3).

Our measurement scheme contains performance indicators that are transparently defined and that support the criterion of reliability. Again, a two-stage procedure was applied by the authors, and a subsample of firms was first independently coded. Then, any uncertainties regarding the definition, interpretation and extrapolation of data were resolved. We ensure content-based overlap between the measurement schemes for the quality of sustainability disclosure and sustainability performance. To precisely assess the theoretical implications derived from voluntary disclosure theory and legitimacy theory, the proxies for sustainability disclosure

[3] For example, a chemical company may report only its total waste weight and be coded as low-quality for reporting element R-E7 because further information about the types of waste and disposal methods is missing. Nevertheless, the information requirement for the corresponding performance indicator P-E4 (total weight of waste) is met.

and sustainability performance must involve and capture similar content, such as by reporting elements regarding emissions and corresponding emissions performance data. Otherwise, changes in sustainability performance are not reflected in what we observe as the quality of sustainability disclosure. The scheme is also multi-dimensional, the performance indicators are material according to the GRI guidelines, and the underlying data refer to directly observable performance values. While these characteristics support the replicability and validity of our measure and mitigate the limitation that we cannot account for every environmental and social impact of the sample firms, our measurement approach is based on the assumption that firms' sustainability disclosure is truthful.[4] We investigate this assumption in section 4.3 based by means of a number of supplemental analyses.

3.3. Sample and methodological approach

The initial sample consists of 388 companies that are included in the Bloomberg European 500 index in January 2013 and are located in France, Germany, Italy, Spain, Sweden, Switzerland or the United Kingdom. Between February and August 2013, English versions of the corporate sustainability reports, if any, the annual report and any web-based sustainability disclosures by the sample firms for reporting year 2011 were identified. The authors manually assessed the sustainability disclosure of each company according to the measurement scheme and recorded the data relevant for evaluating its sustainability performance. Of the companies initially in our sample, 151 were excluded because they had more than two missing performance values in one dimension (or both dimensions). An additional 42 companies were excluded because each industry group is required to have at least five companies to allow for meaningful rescaling of data and comparisons with peers. Table 3 shows the sample selection (Panel A) and the sample distribution by country and industry group (Panel B).

Insert Table 3 about here

The relationship between sustainability disclosure quality and sustainability performance, in addition to the control variables, is assessed by running the following regression models

[4] As long as mandatory sustainability disclosure regulations are missing or remain vague, this caveat applies to all sustainability performance measures applied in previous research. Even sustainability performance measures that are provided by external rating agencies depend to a certain extent on the information provided by the applicable company.

(Clarkson et al., 2008; Clarkson et al., 2011; Dawkins and Fraas, 2011). *HQ_CSD* (equation 1) refers to the level of high-quality corporate sustainability disclosure (hypothesis H1, voluntary disclosure theory), and *LQ_CSD* (equation 2) refers to the level of low-quality corporate sustainability disclosure (hypothesis H2, legitimacy theory).

$$(1)\ HQ\_CSD = \beta_0 + \beta_1\, CSP + \beta_2\, ASSURE + \beta_3\, ORG + \beta_4\, SIZE + \beta_5\, FCF + \beta_6\, LEV + \beta_7\, TOBIN + \sum_{i=8}^{i=13} \beta_i\, COUNTRY + \sum_{j=14}^{j=31} \beta_j\, INDUSTRY + \varepsilon$$

$$(2)\ LQ\_CSD = \beta_0 + \beta_1\, CSP + \beta_2\, ASSURE + \beta_3\, ORG + \beta_4\, SIZE + \beta_5\, FCF + \beta_6\, LEV + \beta_7\, TOBIN + \sum_{i=8}^{i=13} \beta_i\, COUNTRY + \sum_{j=14}^{j=31} \beta_j\, INDUSTRY + \varepsilon$$

Prior empirical studies (Clarkson et al., 2008; Clarkson et al., 2011; Dawkins and Fraas, 2011) rely mainly on one main regression model to test the relationship between sustainability performance and sustainability disclosure as an indicator of the applicability of *either* legitimacy theory *or* voluntary disclosure theory. To test our reasoning that the two theories are not mutually exclusive, we run two regression models. Because a firm can only score in exactly one out of our three disclosure categories for each disclosure item (high-quality, low-quality, or non-disclosure) and both hypotheses are tested based on the same full sample, this methodological approach may entail interdependencies between the *HQ_CSD* and *LQ_CSD* models that are not captured by the non-disclosure category. Such interdependencies fit our theoretical argument that voluntary disclosure theory and legitimacy theory *together* explain firms' reporting behavior. Nevertheless, we address the technical aspect of this concern in section 4.3. of the paper.

3.4. Control variables

We rely on previous research into the determinants of voluntary non-financial disclosure to select our control variables. For instance, Fifka (2013) provides a comprehensive overview. All of the variables are summarized in Table 4 and are described in greater detail below.

Insert Table 4 about here

First, we control for a company's strategic orientation toward sustainability issues, which derives from the strategic management literature (Ullmann, 1985). We draw on external assurance (*ASSURE*) and the hierarchical level of the internal organizational units focused on cor-

porate sustainability (*ORG*) as proxies for strategic orientation. We argue that an active strategic orientation supports the disclosure of credible sustainability-related information, and we thus expect to find a positive (negative) relationship with *HQ_CSD* (*LQ_CSD*). Our variable *ASSURE* indicates whether a company's sustainability disclosure is assured by an external company. Previous research has shown that obtaining external assurance is associated with the strategic integration of sustainability initiatives (Abdel-Khalik, 1993; Knechel et al., 2007). We assume that only companies with an active strategic orientation toward sustainability bear the extra costs of external assurance to demonstrate commitment and credibility. Our variable *ORG* is measured on a four-point rating scale, as shown in Table 4. Previous researchers argue that sustainability reporting requires specific reporting processes and structures (Adams, 2002; Al-Tuwaijri et al., 2004; Ballou et al., 2012; Ruhnke and Gabriel, 2013). The hierarchical level of sustainability-related internal organizational units signals the importance that a company attaches to sustainability (Cowen et al., 1987) and thus indicates the intensity of a firm's strategic orientation toward sustainability.

We also control for firm size (*SIZE*) because many previous empirical studies show a relationship between firm size and sustainability disclosure (Branco and Rodrigues, 2008; Clarkson et al., 2008; Clarkson et al., 2011; Cormier et al., 2005; Dawkins and Fraas, 2011; Kolk, 2003; Patten, 2002). One explanation for this effect focuses on economies of scale with respect to information production costs (Clarkson et al., 2008). Another argument refers to firm size as a proxy for other factors, primarily public visibility (Branco and Rodrigues, 2008; Dawkins and Fraas, 2011) and the extent of monitoring by analysts (Cormier et al., 2005). In either case, a positive relationship between *SIZE* and disclosure quantity is expected. With respect to the quality dimension of sustainability disclosure, an increase in size may lead firms to switch from non-disclosure to low-quality or high-quality disclosure (economies of scale, public visibility, monitoring by analysts), to switch from high-quality to low-quality disclosure (based on legitimacy theory and caution due to outstanding public visibility) or the reverse (stronger demand from analysts). Because those three effects overlap and partially cancel one another out, no sign is expected for the relationship between *SIZE* and *HQ_CSD* or *LQ_CSD*. We measure *SIZE* as the log of the number of employees at the end of the fiscal

year because market capitalization is less stable over time and the book value of total assets or sales is less comparable across industries (e.g., banking and insurance).[5]

We use the free cash flow in millions of euros per employee (*FCF*) at the end of the fiscal year as a proxy for financial performance.[6] One group of researchers (Ullmann, 1985) claims that a company's financial performance determines its financial capacity to invest in and maintain sustainability disclosure. In addition, these researchers argue that only financially sound companies can withstand the (negative) consequences of disclosing proprietary information (Cormier and Magnan, 2003; Cormier et al., 2005). Following this reasoning, the relationship between financial performance and *HQ_CSD* (*LQ_CSD*) should be positive. By contrast, Neu et al. (1998) conclude that companies use environmental disclosures during unprofitable years to demonstrate long-term competitive advantages resulting from environmental investments and posit a negative relationship with financial performance. A third line of research (Patten, 1991) assumes an indifferent relationship between a company's financial performance and its social disclosures and argues that social disclosure is primarily driven by social legitimacy rather than by economic legitimacy. Against the background of these contradictory theoretical considerations and inconsistent empirical results, we do not predict the sign for the relationship between financial performance and *HQ_CSD* or *LQ_CSD*.

We also include the financial leverage of a company (*LEV*) as a proxy for the informational needs of a company's creditors. Financial leverage is measured as a firm's average total assets divided by the firm's average total common equity. It is reasonable to assume that the monitoring demand for information by a company's creditors increases with leverage (Branco and Rodrigues, 2008; Clarkson et al., 2011) and that creditors are interested in a ny's sustainability performance because it may point toward future potential risks related to sustainability issues. Thus, more highly leveraged companies are typically more dependent on creditor demands and accordingly have a greater incentive to inform creditors about their true sustainability performance (Roberts, 1992; Ullmann, 1985). Because high-quality disclo-

[5] We obtain similar (untabulated) results from the regression analyses by employing a log-transformation of market capitalization as an alternative proxy for firm size.

[6] In addition, we re-run the regression analyses with both return on assets (ROA) and return on equity (ROE) as alternative proxies for financial performance, which yield similar (untabulated) results.

sure is assumed to be more reliable and comparable to other disclosure, we expect to find a positive (negative) relationship between financial leverage and *HQ_CSD* (*LQ_CSD*).

Next, we include Tobin's Q (*TOBIN*) in our models as another control variable. Tobin's Q is measured as a firm's market value relative to the replacement cost of its assets and is used to capture information asymmetry (Al-Tuwaijri et al., 2004; Clarkson et al., 2008; Clarkson et al., 2011; Stanny and Ely, 2008). A higher Tobin's Q value reflects a greater degree of information asymmetry. Following the literature on voluntary financial disclosure, information asymmetry between a company's managers and outside investors is assumed to be the major source of demand for financial disclosure (Healy and Palepu, 2001). Consequently, managers strive to decrease information asymmetry through additional disclosure. Because investors prefer high-quality to low-quality information, we expect a positive (negative) relationship between information asymmetry and *HQ_CSD* (*LQ_CSD*).

We also include country and industry group dummies as control variables. A number of empirical studies have revealed systematic, country-specific variations in firms' sustainability disclosure that may arise from regulatory, cultural, or societal differences (Kolk, 2003; Kolk et al., 2001; Orij, 2010; van der Laan Smith et al., 2005). Such differences are closely linked to legitimacy theory because the definition and understanding of legitimacy vary across different countries and societies.[7] Additionally, industry group dummies are included to account for industry-specific effects on the quality of sustainability disclosure. Such industry-specific effects have been demonstrated by a substantial number of previous empirical studies (Cho and Patten, 2007; Dawkins and Fraas, 2011; Patten, 2002; Roberts, 1992).

[7] We examine the country-specific regulatory background for our sample firms for 2011 with respect to the voluntary character of our disclosure items and find no evidence that our measurement scheme for the quality of corporate sustainability disclosure is significantly biased by national regulations. There is only one country – France – that has in place mandatory sustainability disclosure regulations for the sample companies in 2011. However, these French regulations do not necessarily entail high-quality disclosure as defined by our sustainability disclosure measurement scheme.

4. Results

4.1. Descriptive results

Insert Table 5 about here

Panel A of Table 5 presents descriptive statistics for the variables used in the regression analysis. The mean value for *HQ_CSD* is higher than that of *LQ_CSD*, which indicates that, on average, sample firms slightly prefer high-quality to low-quality disclosure. *CSP* varies between 0 and 1, with a mean value slightly above 0.5.

Approximately 68% of the sample firms adopt a sustainability assurance statement, which is a rather high proportion compared with previous studies. For instance, using a panel of Fortune Global 250 firms for 2008, Perego and Kolk (2012) report that 56% of sustainability reports include assurance statements. Because of fundamental differences between the litigation traditions of the European and U.S. markets, the deviation might stem from differences in sample composition, which is restricted to European companies in our study. The deviation may also be attributable to different observation periods because external assurance has become more commonly employed by large companies in recent years (KPMG, 2011). The mean of 2.49 for *ORG* corresponds to a percentage of 81% (untabulated) of the sample firms that run group-wide organizational structures to coordinate their corporate sustainability activities. This percentage exceeds observations from previous studies (e.g., Al-Tuwaijri et al., 2004 report 62%; Ruhnke and Gabriel, 2013 report 59%). However, these deviations likely originate from differing scales and definitions. Because *SIZE* is highly positively skewed, we employ log-transformations to the original data and report the log-transformed data. On average, our sample firms employ 66,037 employees (untabulated) and are thus relatively large companies. Financial performance – as measured by free cash flow on the basis of millions of euros per employee – measures approximately 0.03 on average, and the high standard deviation of financial leverage reflects the diversity of the sample with respect to different industry groups.

Panel B of Table 5 presents the Pearson correlations between all variables in our models. As expected, there is a negative correlation between *HQ_CSD* and *LQ_CSD*, which indicates a close but not perfect relationship between these variables. The correlation coefficient between

CSP and *HQ_CSD* is not significant, whereas the correlation coefficient between *CSP* and *LQ_CSD* is negative. The predicted signs of our control variables are generally in accordance with the correlation statistics, except for *TOBIN*. As expected, there is a positive (negative) relationship between *ASSURE* and *HQ_CSD* (*LQ_CSD*), which indicates that companies with assured sustainability disclosure more often disclose high-quality information. *ORG* is positively correlated with *HQ_CSD*, but the strength of this correlation is weak, and there is no correlation between *ORG* and *LQ_CSD*. Moreover, *ORG* is positively correlated with both *ASSURE* and *SIZE*. The correlation coefficients of both *SIZE* and *FCF* are insignificant, which may reflect the opposing theoretical considerations regarding these control variables. *LEV* is positively (negatively) correlated with *HQ_CSD* (*LQ_CSD*). Counterintuitively, *TOBIN* is negatively (positively) correlated with *HQ_CSD* (*LQ_CSD*), and there is a negative correlation between *ASSURE* and *TOBIN*, which indicates that a lower degree of information asymmetry accompanies more high-quality sustainability disclosure and external assurance, respectively.

4.2. Results of the regression analyses

The results of multivariate regression analyses with robust standard errors (White, 1980) are presented in Table 6. The first set of columns corresponds to hypothesis H1 (*HQ_CSD*), and the second set of columns corresponds to hypothesis H2 (*LQ_CSD*). For each hypothesis, we present three models. Model (a) contains only our main variable of interest, *CSP*, along with *ASSURE* and *ORG*, which capture the strategic orientation toward sustainability issues. In model (b), we include all control variables except country and industry group dummies, whereas model (c) corresponds to our full model.

Insert Table 6 about here

With respect to hypothesis H1, the results of the multivariate regression analysis suggest a positive association between *CSP* and *HQ_CSD* in all three models. This result is consistent with the predictions derived from voluntary disclosure theory that superior sustainability performers disclose sustainability information of high quality rather than of low quality because this type of information is more reliable and comparable. By disclosing primarily high-quality information, these companies actively reveal their superior performance type to the market and are therefore able to distinguish themselves from poor sustainability performers.

The findings for hypothesis H2 reveal a negative relationship between *CSP* and *LQ_CSD* in all three models. The reasoning of legitimacy theory is thus supported, suggesting that poor sustainability performers disclose low-quality rather than high-quality sustainability information to manipulate public perceptions regarding their sustainability performance. Because low-quality information typically lacks reliability and comparability, it is particularly useful for disguising a firm's poor sustainability performance while still contributing to a sustainable company image.

With respect to our control variables, *ASSURE* is significant in all of the models, while SIZE and *LEV* are significant in the reduced models. Thus, in addition to sustainability performance, a company's strategic orientation toward sustainability topics, which is captured by *ASSURE* and *ORG*, is also related to a company's disclosure strategy. Controlling for all other factors, companies that have an active strategic posture more frequently opt for high- rather than low-quality disclosure (models 1b and 1c). On the other hand, a passive strategic orientation toward sustainability issues is associated with predominantly low-quality disclosures (models 2b and 2c). For both high-quality and low-quality sustainability disclosure, the relationship holds for the existence of external control mechanisms (*ASSURE*) and is insignificant for internal organizational structures (*ORG*).

With respect to *SIZE*, there is a positive and significant relationship with *LQ_CSD* in the reduced model but insignificant coefficients for *SIZE* in all of the other models. While this finding indicates that larger firms prefer low-quality sustainability disclosure, we must note that our sample already consists of relatively large firms. Consistent with our expectations, we observe a positive relationship between *LEV* and *HQ_CSD* in the reduced model, indicating that companies with higher leverage are more likely to opt for high-quality sustainability disclosure. We also observe a negative coefficient for *LEV* in the reduced model for low-quality sustainability disclosure, which supports the argument that creditors may be less willing to accept low-quality information and may require higher levels of transparency with an increase in financial leverage. Financial leverage is largely industry-specific, and the change in significance for *LEV* from model (b) to model (c) might stem from the inclusion of industry group dummies. Finally, *FCF* and *TOBIN* are insignificant in all of the models, which is consistent with the prior literature (Clarkson et al., 2008; Clarkson et al., 2011; Dawkins and Fraas, 2011).

Taken all together, the findings from both regression analyses support our position that the two theories are not mutually exclusive but dovetail to simultaneously explain the sustainability reporting behavior of our sample firms. In the next section, we provide a battery of robustness checks to test whether our results are sensitive to our sustainability disclosure quality and sustainability performance measurements.

4.3. Supplemental analyses

The truthfulness of the voluntarily disclosed information is an important assumption of our research design that concerns the measurement of both sustainability disclosure quality and sustainability performance. An effective mechanism to ensure the correctness and accuracy of reported sustainability information is external assurance (O'Dwyer, 2011). Considering the high proportion of external assurance among our sample firms – 68% obtain external assurance regarding their sustainability disclosure – untruthful disclosure does not appear to be a major problem for our study. Nonetheless, we perform several additional analyses to check the sensitivity of our findings with respect to the assumption regarding truthful disclosure.

Based on the assumption that the probability of untruthful disclosure is presumably higher among firms without external assurance than among firms that obtain external assurance, we re-run the regression analyses for the subsamples of firms with external assurance (n=133) and without external assurance (n=62) separately. If our main results are biased by untruthful disclosure, we would expect to obtain different results for the subsample of companies without external assurance. However, the results from the re-estimation of the regression analyses (omitting ASSURE) clearly support both hypotheses within each subsample and thereby suggest that there are no concerns regarding the truthful disclosure assumption. Nonetheless, we cannot completely rule out that untruthful disclosure occurs in the subsample of firms *without* external assurance (n=62), and we thus apply two methodologies to further investigate the truthfulness of sustainability disclosure for this subsample.

First, we investigate whether the 14 disclosure items in our measurement scheme are subject to restatements in the subsequent reporting year. In total, we identify five reports that include restatements of the 2011 data for at least one of the 14 disclosure items. Three such reports contain restatements of one indicator, and the remaining two reports contain restatements of

two and three indicators, respectively. Of the eight restated indicators, three indicators have been restated "positively" (i.e., revealing better performance data ex post), whereas five indicators have been restated "negatively" (i.e., revealing worse performance data ex post). The primary reasons for restatements are changes in the methodology related to data collection. Only one company reports errors in 2011, and one company does not specify the reasons for the restatement. Taken together, the findings from this additional analysis indicate no concerns with respect to assumptions about truthful disclosure.

Second, we use data from the RepRisk ESG Risk Platform to check whether the sustainability disclosure of firms without external assurance has been subject to criticism from third parties since 2007. RepRisk captures and analyzes information based on a rules-based and systematic methodology of screening and monitoring over 80,000 media, stakeholder, and other third-party sources external to the company on a global scale with respect to environmental, social, and governance (ESG)-related risk incidents (RepRisk, 2016). The database, launched in 2007, contains data that is updated daily on over 65,000 public and private companies from around the world that have been exposed to ESG-related risk incidents. For each of the 62 companies, we check the database for criticism regarding the firms' sustainability disclosure based on the search terms "misleading communication" (a pre-defined category), "disclosure", "reporting", "greenwash", "false", and "erroneous". We manually assess the list of the results. Most entries refer to criticism due to misleading advertisement and only one risk incident directly relates to firms' sustainability disclosure.[8] Overall, the results from this third-party reporting search do not reveal any evidence of untruthful disclosure that might bias our measures of both the quality of sustainability disclosure and sustainability performance.

Next, we check the robustness of our sustainability performance measure with respect to various aspects. For the subsample of firms with external assurance, we manually assess the extent to which these external assurances cover the eight sustainability performance indicators in our measurement scheme. On average, 80% of the sustainability performance indicators are

[8] This news entry recounts a report on flaws in the GRI application levels and disclosure categories oft he GRI content index by European electricity companies. Because our measurement approach to the quality of sustainability disclosure relies on a content-based analysis of the respective disclosure items, such a confounding reporting practice does not affect the quality measure of our sustainability disclosure.

explicitly subject to an external assurance, which provides additional confidence in the reliability of the data. We also experiment with different winsorization levels and re-run our models with sustainability performance data winsorized at the 5% and 1% levels (rather than at the 10% level, as in our baseline model). The results are similar to those of the initial model and are thus not significantly influenced by how we handle spurious outliers (results untabulated).

Because we propose a measurement of sustainability performance in this study that has not yet been established in the literature, we also examine whether an alternative measurement yields similar results. Similar to Dhaliwal et al. (2011), we re-run the regression models for both *HQ_CSD* (equation 1) and *LQ_CSD* (equation 2) by using membership in the Dow Jones Sustainability Index (DJSI) as a proxy for sustainability performance (S&P Dow Jones Indices and RobecoSAM, 2014). The variable *DJSI* assumes the value of one if a company belongs to the DJSI Europe in 2011 and zero otherwise. The results from the regression analyses using *DJSI* instead of *CSP* as the main variable of interest support our initial findings. With respect to hypothesis H1 (hypothesis H2), all models yield a positive (negative) association between *DJSI* and *HQ_CSD* (*LQ_CSD*) (results untabulated).

Third, we address potential concerns regarding the independence of the coding of each disclosure item as high-quality, low-quality, or non-disclosure. Due to the limited use of non-disclosure by our sample firms, it might be argued that separate testing for each hypothesis based on the same dataset creates empirical results for each hypothesis that are a replication of one another with reversed correlations. To eliminate possible links between the results for our *HQ_CSD* (equation 1) and *LQ_CSD* (equation 2) models, we separately test the models based on randomly drawn subsamples that provides for no overlap. The full sample of 195 firms is randomly split into the 98 firms that are used to test hypothesis H1 (*HQ_CSD* model) and the 97 firms that are used to test hypothesis H2 (*LQ_CSD* model). This procedure is repeated 10,000 times. Figure 1 presents the regression coefficients for *CSP* together with their corresponding p-values for equation (1) (left side) and equation (2) (right side).

Insert Figure 1 about here

The lower sizes of the subsample considerably reduce the statistical power of the regression models. Nevertheless, we obtain positive *CSP* coefficient estimates b_{CSP} for the *HQ_CSD*

model (equation 1) that are significant at the 10% level or higher in 71.85% of all cases (left graph, right lower quadrant); with respect to the *LQ_CSD* model (equation 2), we obtain negative estimates that are significant at the 10% level or higher in 66.28% of all cases (right graph, right lower quadrant). Another 27.57% (32.63%) of the *CSP* estimates have the expected positive (negative) sign but are not statistically significant. These findings attenuate potential concerns with respect to our research setting and support our initial findings regarding the applicability of both voluntary disclosure theory and legitimacy theory to explain the reporting behavior of our sample firms.

Finally, we account for the problem of missing data with regard to certain performance indicators in the sample. According to our measurement scheme for sustainability performance, these missing values are replaced with the mean values of the remaining indicators for each dimension. To assess the robustness of our results with respect to missing values, we cover a wide range of possible scenarios using the Monte Carlo simulation method (Metropolis and Ulam, 1949). Beginning with winsorized original data, missing values are replaced by simulated values randomly drawn from a uniform distribution and separately rescaled for each industry group within an interval of [0.8*minimum per industry; 1.2*maximum per industry]. The additional range of 20% accounts for the possibility that missing data constitute unobserved extremes in our sample. All subsequent data rescaling and aggregation procedures are identical to our original method. The new dataset containing observed and simulated values is used to run our main regression models to test hypothesis 1 (hypothesis 2). These steps are repeated 100,000 times, and the results are presented in Figure 2.

Insert Figure 2 about here

Each mark represents a simulated regression coefficient b_{CSP} and the corresponding p-value for *HQ_CSD* (left side) and *LQ_CSD* (right side). For all simulated scenarios, the signs of the regression coefficients for *CSP* remain positive (negative). With respect to *HQ_CSD*, 98.02% of all scenarios yield statistically significant results at the 1% level (1.97% at the 5% level and 0.01% at the 10% level). There is no scenario with insignificant results, which is indicated by the broken line in Figure 2. For *LQ_CSD*, 82.18% of all scenarios are significant at the 1% level (17.54% at the 5% level and 0.28% at the 10% level). In this case, only 0.01% of the results from all scenarios are insignificant (marks above the broken line). Overall, the results

of the Monte Carlo simulation comprehensively support the robustness of our findings with respect to missing performance values in our data.

5. Conclusions

Voluntary disclosure theory and legitimacy theory are the prevailing theoretical foundational concepts used in the literature to explain the relationship between sustainability performance and sustainability disclosure. However, empirical researchers typically regard these two theories as incompatible with one another – even mutually exclusive – and interpret evidence supporting one of the theories as evidence disproving the other. Against the background of mixed empirical evidence, some researchers have recently revised the conjecture that these two theories are mutually exclusive and called for a switch in the "focus of enquiry" (Clarkson et al., 2008).

We respond to this call and present theoretical reasoning and empirical evidence that reconciles the two theories by redirecting the focus of inquiry from the quantity of corporate sustainability disclosure to its quality. In accordance with voluntary disclosure theory, we argue that superior sustainability performers prefer high-quality sustainability disclosure because it is more transparent, reliable and comparable. In addition, we build on legitimacy theory and predict a negative relationship between sustainability performance and low-quality sustainability disclosure because poor sustainability performers avoid transparency to protect their image as sustainable firms. The results from regression analysis performed on a sample of 195 European companies support this reasoning and reveal a positive (negative) and significant relationship between sustainability performance and high-quality (low-quality) sustainability disclosure. The results are robust to a number of additional analyses and robustness checks. Moreover, we respond to calls from several researchers and present improved measurement approaches for both sustainability performance and sustainability disclosure quality.

As is typical, the results of this paper are also subject to certain limitations. First, the generalizability of our findings depends on both our sample and on the time period of our study. Our sample refers to the 2011 reporting period and predominantly consists of large and publicly listed companies. Therefore, our results may not hold for other periods, for small firms, or for firms less oriented to the capital markets. A second caveat of our study is linked to the truthfulness of disclosure, an important assumption of our research design. However, we perform a

number of additional analyses to check the robustness of our results with respect to this assumption. None of these results indicate any concerns with respect to the truthfulness of disclosure. Another caveat applies to the problem of non-reporting companies. Our sustainability performance measurement scheme requires that a company reveals sufficient performance data to be evaluated; as a consequence, non-reporting firms are excluded from the sample, which indicates that our results cannot account for those firms. Biases may also arise from companies that were included and partly withheld data. Nevertheless, the results from a Monte Carlo simulation of these missing performance values do not indicate reasonable concerns with respect to the robustness of our findings resulting from incomplete performance data of our sample firms.

In addition to the contributions to the academic literature, our study also has practical implications that may lead to future research. The finding that superior sustainability performers use high-quality sustainability disclosure to signal their sustainability performance to the market, whereas poor sustainability performers use low-quality sustainability disclosure to attempt to positively influence public perceptions, may point toward the need for a precise and binding regulatory framework for the contents of sustainability reports. However, there is empirical evidence indicating that firms' compliance with such mandatory sustainability disclosure regulations is often low (Chauvey et al., 2015; Larrinaga et al., 2002). Future research could therefore investigate different types of regulation of sustainability disclosure and analyze under which conditions mandatory sustainability disclosure regulations can achieve high-quality sustainability disclosure. In this respect, the introduction of mandatory sustainability reporting by the European Union (Directive 2014/95/EU) yields an interesting research setting. Future research might investigate both the pre-regulation adaptations of reporting behavior and post-regulation sustainability disclosure quality to determine the effectiveness of the new regulatory frameworks.

Second, the results of our study provide preliminary evidence regarding the relevance of high-quality sustainability disclosure for capital market participants. A different research design is necessary to test whether high-quality sustainability disclosure is indeed appraised by capital

market participants and whether it affects firm value.[9] Although the results from previous investigations on the value relevance of sustainability disclosure in general are promising (Clarkson et al., 2013; Dhaliwal et al., 2012), the integration of the quality dimension of sustainability disclosure would add a new perspective to the ongoing discussion in this field of research.

[9] This research question must be separated from the overwhelming number of investigations into the relationship between sustainability performance and financial performance (for an overview, see Dixon-Fowler et al., 2013; Orlitzky et al., 2003).

Table 1. Measurement scheme for corporate sustainability disclosure quality

Code	Disclosure item	Minimum requirements for high-quality score	GRI G3.1
Environmental dimension			
R-E1	materials used	all substantial input materials by weight or volume[a]	EN1
R-E2	energy consumption and renewables	direct and indirect energy consumption, share of renewable energy sources[b]	EN3/4
R-E3	water withdrawal	total withdrawal by source[c]	EN8
R-E4	greenhouse gas emissions	total direct and indirect emissions (GHG protocol scopes 1, 2, and 3)	EN16/17
R-E5	ozone-depleting substances and other air emissions	total emissions of ozone-depleting substances; other significant air emissions by type and weight for at least one substance; alternatively, an explicit statement of irrelevance for both[d]	EN19/20
R-E6	water discharge	total discharge by quality (emissions to water by type and weight for at least one substance; alternatively, an explicit statement of irrelevance) and destination[e]	EN21
R-E7	waste	total weight by type and disposal method[f] *maximum environmental score is 7*	EN22
Social dimension			
R-S1	workforce	total workforce based on at least three criteria (division, region, employment type, employment contract, qualification, age or gender)	LA1
R-S2	employee turnover	total number of employees leaving for any reason (not for a single reason only)	LA2
R-S3	collective bargaining agreements	percentage of total workforce covered by collective bargaining agreements	LA4
R-S4	safety and health	work safety and health based on at least two criteria (rates of injury, occupational diseases, lost days, absenteeism, fatalities)	LA7
R-S5	training	training (or related) time	LA10
R-S6	discrimination	total number of incidents or explicit statement that no incidents occurred (no indirect paraphrasing or references to codes of conduct)	HR4
R-S7	child, forced, and compulsory labor	scope and numerical results of audits (within company or supply chain) regarding at least one aspect *maximum social score is 7*	HR6/7

[a] Adjustments for industry groups 13-15 (see description in Table 3): use of paper is sufficient.
[b] Adjustments for industry group 12 (see description in Panel B of Table 3): share of renewable energy sources is excluded.
Adjustments for industry group 18 (see description in Panel B of Table 3): share of renewable energy *produced*.
[c] Adjustments for industry groups 4-6 and 13-15 (see description in Panel B of Table 3): by source is excluded.
[d] Adjustments for industry groups 4-6 and 13-15 (see description in Panel B of Table 3): ozone-depleting substances *or* other significant air emissions.
[e] Adjustments for industry groups 4-6 and 13-15 (see description in Panel B of Table 3): by quality and destination is excluded.
[f] Adjustments for industry groups 13-15 (see description in Panel B of Table 3): by type and disposal method is excluded.

This table presents the indicators for the assessment of sustainability disclosure quality with respect to the environmental and social, in particular employee-related, reporting dimensions. The first two columns contain the respective disclosure items. Column three presents the minimum requirements that are defined for high-quality disclosure. The fourth column indicates the link to the G3.1 guidelines (GRI, 2011b, 2011c, 2011d).

Table 2. Measurement scheme for corporate sustainability performance

Code	Performance indicator	Measurement	Unit	Link sustainability disclosure quality (Table 1)
Environmental dimension				
P-E1	energy consumption	(direct + indirect energy consumption) / number of employees[a]	MWh / employee	R-E2
P-E2	water withdrawal	(total water withdrawal – cooling water) / number of employees[a]	m^3 / employee	R-E3
P-E3	greenhouse gas emissions	(GHG protocol scope 1 + scope 2 emissions) / number of employees[a]	t / employee	R-E4
P-E4	total weight of waste	total weight of waste / number of employees[a]	t / employee	R-E7
Social dimension				
P-S1	employee turnover	total number of employees[a] who leave / number of employees[a] * 100	percent	R-S2
P-S2	lost time incident rate	number of incidents resulting in lost time from work / (total hours worked / 200,000)	incidents / h	R-S4
P-S3	employee training	total training time / number of employees[a]	h / employee	R-S5
P-S4	share of women in the highest corporate bodies[2]	total number of women in the highest corporate bodies[b] / total number of members of the highest corporate bodies[b] * 100	percent	R-S6

[a] If available, full-time equivalents; headcount otherwise.
[b] Management board and supervisory board for two-tier system countries; board of directors for one-tier system countries.

This table presents the indicators for the measurement of corporate sustainability performance with respect to the environmental and social, in particular employee-related, dimensions. Columns one and two contain indicator names and brief descriptions. Columns three and four present formulas for calculations and the units of measurement. The last column indicates how the sustainability performance measurement is linked to the measurement of sustainability disclosure quality in Table 1.

Table 3. Sample selection and distribution

Panel A: Sample selection	
Initial Bloomberg European 500 sample	500
Less: Firms not domiciled in France, Germany, Italy, Spain, Sweden, Switzerland or the United Kingdom	−112
	388
Less: Firms with insufficient data on sustainability performance as described in section 3.3	−151
	237
Less: Firms that belong to an industry group with less than five companies	−42
Total sample size	**195**

Panel B: Sample distribution			
By country	**n**	**By industry group**	**n**
1 France	44	1 chemicals	10
2 Germany	28	2 building materials, paper, steel	16
3 Italy	16	3 mining	11
4 Spain	27	4 advertising, entertainment, media	11
5 Sweden	13	5 telecommunication	11
6 Switzerland	14	6 computers and software	5
7 United Kingdom	53	7 car manufacturers	8
		8 retail	5
		9 food and beverages	9
		10 commercial services	7
		11 pharmaceuticals	8
		12 oil and gas	8
		13 banking and financial services	30
		14 insurance	10
		15 REITS	10
		16 aerospace and defense	8
		17 engineering and construction	11
		18 electricity	7
		19 gas, water, and electricity distribution	10
Total	**195**	**Total**	**195**

Table 4. Variables overview

	HQ_CSD **predicted sign**	***LQ_CSD*** **predicted sign**	**Description**	**Source of data**
HQ_CSD			high-quality sustainability disclosure score for reporting year 2011, measured as described in Table 1 and in section 3.1	manual
LQ_CSD			low-quality sustainability disclosure score for reporting year 2011, measured as described in Table 1 and in section 3.1	manual
CSP	+	−	level of sustainability performance for reporting year 2011, measured as described in Table 2 and in section 3.2	manual
ASSURE	+	−	external assurance, equals "1" if a company's sustainability disclosure for reporting year 2011 is assured by an external company and "0" otherwise	manual
ORG	+	−	highest hierarchical level of internal organizational units focused on corporate sustainability during the 2011 reporting year, measured on a four-point rating scale: (1) a company's corporate sustainability activities are supervised by only a few local standalone units with no superordinate organizational unit on the group level (2) there is a specialized organizational unit determining and supervising the entire group's corporate sustainability strategy (3) at least one member of this specialized organizational unit is a representative of the highest corporate bodies (4) a main committee of the board of directors is dedicated to the firm's corporate sustainability activities	manual
SIZE			firm size, measured as the log of total employees at the end of fiscal year 2011 (full-time equivalents if available, headcount otherwise)	manual
FCF			financial performance, measured as the free cash flow (cash flow from operating activities − total capital expenditures) in millions of euros per employee at the end of fiscal year 2011 (full-time equivalents if available, headcount otherwise)	Bloomberg
LEV	+	−	financial leverage (average total assets / average total common equity)	Bloomberg
TOBIN	+	−	Tobin's Q, measured as (market value common equity + book value long-term debt and current liabilities) / book value total assets at the end of fiscal year 2011	Bloomberg
COUNTRY			country of domicile dummy variables as reported in Panel B of Table 3, and the reference category is the UK	Bloomberg
INDUSTRY			industry group dummy variables as reported in Panel B of Table 3, and the reference category is banking and financial services	Bloomberg/ manual

Table 5. Descriptive and correlation statistics for regression variables

Panel A: Descriptive statistics

	n	mean	Median	sd	25th percentile	75th percentile
(1) *HQ_CSD*	195	6.98	7.00	2.88	5.00	9.00
(2) *LQ_CSD*	195	5.38	5.00	2.32	4.00	7.00
(3) *CSP*	195	0.55	0.57	0.19	0.42	0.70
(4) *ASSURE*	195	0.68	1.00	0.47	0.00	1.00
(5) *ORG*	195	2.49	2.00	1.10	2.00	4.00
(6) *SIZE*	195	4.42	4.61	0.76	4.01	4.97
(7) *FCF*	195	0.03	0.02	0.32	0.00	0.06
(8) *LEV*	195	6.91	3.40	11.52	2.27	6.82
(9) *TOBIN*	195	1.25	1.06	0.48	0.98	1.30

Panel B: Correlation statistics

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
(1)	1.0000								
(2)	−0.8042 (0.0000)	1.0000							
(3)	0.0974 (0.1757)	−0.1187 (0.0983)	1.0000						
(4)	0.3991 (0.0000)	−0.2165 (0.0004)	−0.1004 (0.1624)	1.0000					
(5)	0.1319 (0.0660)	0.0202 (0.7787)	−0.0143 (0.8423)	0.3151 (0.0000)	1.0000				
(6)	0.0753 (0.2191)	0.1326 (0.0300)	0.0612 (0.3956)	0.2317 (0.0001)	0.2370 (0.0009)	1.0000			
(7)	0.0525 (0.3923)	−0.0458 (0.4556)	0.1061 (0.1398)	−0.0424 (0.4894)	−0.0575 (0.4248)	0.2150 (0.0004)	1.0000		
(8)	0.1439 (0.0184)	−0.2227 (0.0002)	0.0014 (0.9850)	−0.0218 (0.7219)	−0.1345 (0.0608)	0.0109 (0.8596)	0.1608 (0.0083)	1.0000	
(9)	−0.1394 (0.0225)	0.1349 (0.0273)	0.0779 (0.2791)	−0.1263 (0.0387)	0.0440 (0.5414)	−0.0695 (0.2570)	0.0481 (0.4332)	−0.2492 (0.0000)	1.0000

(1) *HQ_CSD*, (2) *LQ_CSD*, (3) *CSP*, (4) *ASSURE*, (5) *ORG*, (6) *SIZE*, (7) *FCF*, (8) *LEV*, (9) *TOBIN*

Table 6 presents descriptive statistics (Panel A) and correlation statistics (Panel B) for the variables used in the regression analysis. Statistics are presented for the full sample of 195 firms. Panel B reports bivariate Pearson correlation coefficients and p-values (in parentheses) for a two-tailed test of statistical significance. *HQ_CSD* refers to high-quality disclosure, *LQ_CSD* refers to low-quality disclosure, and *CSP* proxies for corporate sustainability performance. *ASSURE* is a dummy variable indicating external assurance, *ORG* proxies for the organizational integration of sustainability, and *SIZE* and *FCF* refer to company size and financial performance, respectively. *LEV* refers to financial leverage, and *TOBIN* proxies for information asymmetry. Detailed descriptions of all variables are provided in sections 3.1., 3.2., and 3.4.

Table 6. Regression results

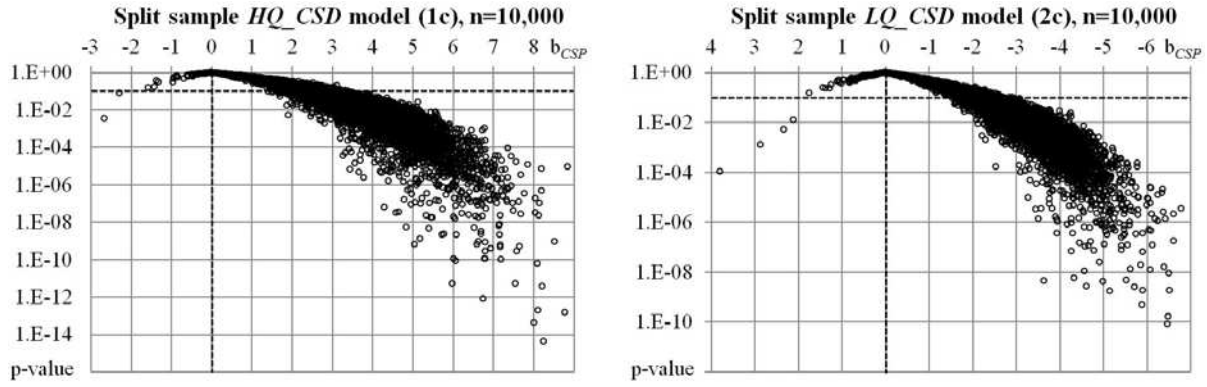
(1) $HQ\_CSD = \beta_0 + \beta_1\ CSP + \beta_2\ ASSURE + \beta_3\ ORG + \beta_4\ SIZE + \beta_5\ FCF + \beta_6\ LEV + \beta_7\ TOBIN + \sum_{i=8}^{i=13} \beta_i\ COUNTRY + \sum_{j=14}^{j=31} \beta_j\ INDUSTRY + \varepsilon$

(2) $LQ\_CSD = \beta_0 + \beta_1\ CSP + \beta_2\ ASSURE + \beta_3\ ORG + \beta_4\ SIZE + \beta_5\ FCF + \beta_6\ LEV + \beta_7\ TOBIN + \sum_{i=8}^{i=13} \beta_i\ COUNTRY + \sum_{j=14}^{j=31} \beta_j\ INDUSTRY + \varepsilon$

	(1) *HQ_CSD* models			**(2) *LQ_CSD* models**		
	(1a)	**(1b)**	**(1c)**	**(2a)**	**(2b)**	**(2c)**
Intercept	3.6605***	3.7827***	3.5434*	6.7406***	5.2248***	5.0303***
	(4.6474)	(2.6171)	(1.8771)	(9.8235)	(4.3195)	(3.2326)
CSP	2.6614**	2.7335**	3.3036***	−1.9094**	−2.0598**	−2.6497***
	(2.5332)	(2.5697)	(3.3144)	(−2.0862)	(−2.3139)	(−3.2248)
ASSURE	2.5350***	2.6020***	2.1922***	−1.2620***	−1.4574***	−1.1729***
	(5.7187)	(5.7167)	(4.9277)	(−3.2680)	(−3.8263)	(−3.1983)
ORG	0.0533	0.1401	0.2045	0.2051	0.0593	−0.0388
	(0.2859)	(0.7311)	(1.0171)	(1.2627)	(0.3697)	(−0.2340)
SIZE		−0.1397	−0.1051		0.5887**	0.4521
		(−0.4964)	(−0.2986)		(2.4989)	(1.5580)
FCF		0.5415	−0.4773		−0.2333	0.1162
		(0.8336)	(−0.6984)		(−0.4291)	(0.2062)
LEV		0.0409**	0.0141		−0.0622***	−0.0199
		(2.2799)	(0.6805)		(−4.1477)	(−1.1681)
TOBIN		−0.0737	0.1696		−0.0396	−0.3283
		(−0.1739)	(0.3690)		(−0.1117)	(−0.8667)
Country dummies	NO	NO	YES	NO	NO	YES
Industry group dummies	NO	NO	YES	NO	NO	YES
Observations	195	195	195	195	195	195
Adjusted R^2	0.1648	0.1770	0.3635	0.0530	0.1407	0.3349
F-Statistic	13.7628***	6.96***	4.5747***	4.62***	5.5382***	4.1511***

This table reports ordinary least squares coefficient estimates and t-statistics (in parentheses) based on Huber-White robust standard errors. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels (two-tailed), respectively. *HQ_CSD* refers to high-quality disclosure, *LQ_CSD* refers to low-quality disclosure, and *CSP* proxies for corporate sustainability performance. *ASSURE* is a dummy variable indicating external assurance, *ORG* proxies for the organizational integration of sustainability, and *SIZE* and *FCF* refer to company size and financial performance, respectively. *LEV* refers to financial leverage, and *TOBIN* proxies for information asymmetry. Detailed descriptions of all variables are provided in sections 3.1., 3.2., and 3.4. *COUNTRY* and *INDUSTRY* refer to dummy variables for the seven countries of domicile and 19 industry groups that are summarized in Table 3.

Figure 1. Results from the simulation of separated hypotheses testing



This figure shows *CSP* coefficient estimates b_{CSP} and corresponding p-values when our hypotheses are separately tested using non-overlapping subsamples. The left graph corresponds to the *HQ_CSD* model (equation 1) and shows 10,000 regression results for different randomly drawn subsamples (n=98) out of our full sample (n=195). The right side shows the results for the *LQ_CSD* model (equation 2), which are obtained from the remaining subsample (n=97). The broken horizontal lines illustrate the 10% statistical significance level, and results below these lines refer to higher levels of statistical significance.

Figure 2. Results from the simulation of missing performance values

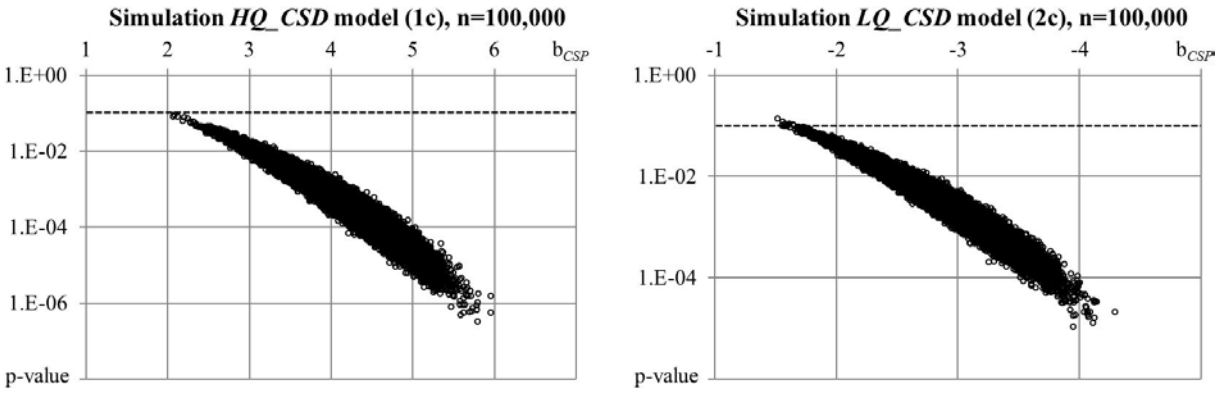


Figure 2 shows *CSP* coefficient estimates b_{CSP} and corresponding p-values when missing performance values are replaced by simulated performance values. The left (right) graph corresponds to the *HQ_CSD* (*LQ_CSD*) model, i.e., equation 1 (equation 2), and shows 100,000 regression results for different datasets that contain all original data plus different randomly simulated values for missing data. The broken horizontal lines illustrate the 10% statistical significance level, and results below these lines refer to higher levels of statistical significance.

REFERENCES

Abdel-Khalik, A.R., 1993. Why do private companies demand auditing? A case for organizational loss of control. J Account Audit Financ 8(1), 31-52.

Adams, C.A., 2002. Internal organisational factors influencing corporate social and ethical reporting: Beyond current theorising. Account Audit Accountab J 15(2), 223-250.

Aerts, W., Cormier, D., 2009. Media legitimacy and corporate environmental communication. Account Org Soc 34(1), 1-27.

Aerts, W., Cormier, D., Magnan, M., 2008. Corporate environmental disclosure, financial markets and the media: An international perspective. Ecol. Econ. 64(3), 643-659.

Al-Tuwaijri, S.A., Christensen, T.E., Hughes, K.E., 2004. The relations among environmental disclosure, environmental performance, and economic performance: A simultaneous equations approach. Account Org Soc 29(5-6), 447-471.

Ameer, R., Othman, R., 2012. Sustainability practices and corporate financial performance: A study based on the top global corporations. J Bus Ethics 108(1), 61-79.

Ballou, B., Casey, R.J., Grenier, J.H., Heitger, D.L., 2012. Exploring the strategic integration of sustainability initiatives: Opportunities for accounting research. Account Horiz 26(2), 265-288.

Ballou, B., Heitger, D.L., Landes, C.E., 2006. The future of corporate sustainability reporting. J Accountancy 202(6), 65-74.

Bewley, K., Li, Y., 2000. Disclosure of environmental information by Canadian manufacturing companies: A voluntary disclosure perspective. Adv Env Account Manage 1, 201-226.

Beyer, A., Cohen, D.A., Lys, T.Z., Walther, B.R., 2010. The financial reporting environment: Review of the recent literature. J Account Econ 50, 296-343.

Branco, M.C., Rodrigues, L.L., 2008. Factors influencing social responsibility disclosure by Portuguese companies. J Bus Ethics 83(4), 685-701.

Chatterji, A.K., Levine, D.I., Toffel, M.W., 2009. How well do social ratings actually measure corporate social responsibility? J. Econ. Manage. Strat. 18(1), 125-169.

Chauvey, J.N., Giordano-Spring, S., Cho, C.H., Patten, D.M., 2015. The normativity and legitimacy of CSR disclosure: Evidence form France. J Bus Ethics 130(4), 789-803.

Chin, M.K., Hambrick, D.C., Treviño, L.K., 2013. Political ideologies of CEOs: The influence of executives' values on corporate social responsibility. Admin Sci Quart 58(2), 197-232.

Chiu, S.C., Sharfman, M., 2011. Legitimacy, visibility, and the antecedents of corporate social performance: An investigation of the instrumental perspective. J. Manage. 37(6), 1558-1585.

Cho, C.H., Guidry, R.P., Hageman, A.M., Patten, D.M., 2012. Do actions speak louder than words? An empirical investigation of corporate environmental reputation. Account Org Soc 37(1), 14-25.

Cho, C.H., Patten, D.M., 2007. The role of environmental disclosures as tools of legitimacy: A research note. Account Org Soc 32(7-8), 639-647.

Cho, C.H., Patten, D.M., Roberts, R.W., 2006. Corporate political strategy: An examination of the relation between political expenditures, environmental performance, and environmental disclosure. J Bus Ethics 67(2), 139-154.

Clarkson, P.M., Fang, X., Li, Y., Richardson, G.D., 2013. The relevance of environmental disclosures: Are such disclosures incrementally informative? J Account Public Pol 32(5), 410-431.

Clarkson, P.M., Li, Y., Richardson, G.D., Vasvari, F.P., 2008. Revisiting the relation between environmental performance and environmental disclosure: An empirical analysis. Account Org Soc 33(4-5), 303-327.

Clarkson, P.M., Overell, M.B., Chapple, L., 2011. Environmental reporting and its relation to corporate environmental performance. Abacus 47(1), 27-60.

Cormier, D., Magnan, M., 2003. Environmental reporting management: A continental European perspective. J Account Public Pol 22(1), 43-62.

Cormier, D., Magnan, M., van Velthoven, B., 2005. Environmental disclosure quality in large German companies: Economic incentives, public pressures or institutional conditions? Eur. Account. Rev. 14(1), 3-39.

Cowen, S.S., Ferreri, L.B., Parker, L.D., 1987. The impact of corporate characteristics on social responsibility disclosure: A typology and frequency-based analysis. Account Org Soc 12(2), 111-122.

Davis, K., 1973. The case for and against business assumption of social responsibilities. Acad Manage J 16(2), 312-322.

Dawkins, C., Fraas, J.W., 2011. Coming clean: The impact of environmental performance and visibility on corporate climate change disclosure. J Bus Ethics 100(2), 303-322.

De Villiers, C., Marques, A., 2016. Corporate social responsibility, country-level predispositions, and the consequences of choosing a level of disclosure. Account. Bus. Res. 46(2), 167-195.

de Villiers, C., van Staden, C.J., 2006. Can less environmental disclosure have a legitimising effect? Evidence from Africa. Account Org Soc 31(8), 763-781.

Deegan, C., 2002. Introduction: The legitimizing effect of social and environmental disclosures – A theoretical foundation. Account Audit Accountab J 15(3), 282-311.

Dhaliwal, D.S., Li, O.Z., Tsang, A., Yang, Y.G., 2011. Voluntary nonfinancial disclosure and the cost of equity capital: The initiation of corporate social responsibility reporting. Account Rev 86(1), 59-100.

Dhaliwal, D.S., Radhakrishnan, S., Tsang, A., Yang, Y.G., 2012. Nonfinancial disclosure and analyst forecast accuracy: International evidence on corporate social responsibility disclosure. Account Rev 87(3), 723-759.

Dixon-Fowler, H.R., Slater, D.J., Johnson, J.L., Ellstrand, A.E., Romi, A.M., 2013. Beyond "does it pay to be green?" A meta-analysis of moderators of the CEP–CFP relationship. J Bus Ethics 112(2), 353-366.

Elkington, J., 1997. Cannibals with forks: The triple bottom line of 21st century business. Capstone, Oxford.

Fifka, M.S., 2013. Corporate responsibility reporting and its determinants in comparative perspective – A review of the empirical literature and a meta-analysis. Bus. Strat. Env. 22(1), 1-35.

Freeman, R.E., 1984. Strategic management: A stakeholder approach. Cambridge University Press, Boston, MA.

Gray, R., 2006. Social, environmental and sustainability reporting and organisational value creation? Whose value? Whose creation? Accounting, Auditing and Accountability Journal 19(6), 793-819.

GRI, 2011a. Sustainability reporting guidelines version 3.1. GRI, The Netherlands.

GRI, 2011b. Indicator protocols set environment (EN). GRI, The Netherlands.

GRI, 2011c. Indicator protocols set labor practices and decent work (LA). GRI, The Netherlands.

GRI, 2011d. Indicator protocols set human rights (HR). GRI, The Netherlands.

Healy, P.M., Palepu, K.G., 2001. Information asymmetry, corporate disclosure, and the capital markets: A review of the empirical disclosure literature. J Account Econ 31(1-3), 405-440.

Hong, Y., Andersen, M.L., 2011. The relationship between corporate social responsibility and earnings management: An exploratory study. J Bus Ethics 104(4), 461-471.

Horváthová, E., 2012. The impact of environmental performance on firm performance: Short-term costs and long-term benefits? Ecol. Econ. 84, 91-97.

Hughes, S.B., Anderson, A., Golden, S., 2001. Corporate environmental disclosures: Are they useful in determining environmental performance? J Account Public Pol 20(3), 217-240.

Knechel, W.R., Salterio, S.E., Ballou, B., 2007. Auditing: Assurance and risk, 3rd ed. Thomson/South-Western, Mason, OH.

Kolk, A., 2003. Trends in sustainability reporting by the Fortune Global 250. Bus. Strat. Env. 12(5), 279-291.

Kolk, A., Walhain, S., van de Wateringen, S., 2001. Environmental reporting by the Fortune Global 250: Exploring the influence of nationality and sector. Bus. Strat. Env. 10(1), 15-28.

KPMG, 2011. KPMG international survey of corporate responsibility reporting 2011. KPMG, Switzerland.

Larrinaga, C., Carrasco, F., Correa, C., Llena, F., Moneva, j.M., 2002. Accountability and accounting regulation: the case of the Spanish environmental disclosure standard. Eur. Account. Rev. 11(4), 723-740.

Leuz, C., Wysocki, P., 2008. Economic consequences of financial reporting and disclosure regulation: a review and suggestions for future research. http:/ssrn.com/abstract=1105398.

Li, Y., Richardson, G.D., Thornton, D.B., 1997. Corporate disclosure of environmental liability information: Theory and evidence. Contemp Account Res 14(3), 435-474.

Metropolis, N., Ulam, S., 1949. The Monte Carlo method. J. Am. Stat. Assoc. 44(247), 335-341.

Neu, D., Warsame, H., Pedwell, K., 1998. Managing public impressions: Environmental disclosures in annual reports. Account Org Soc 23(3), 265-282.

O'Dwyer, B., 2011. The case of sustainability assurance: Constructing a new assurance service. Contemp Account Res 28(4), 1230-1266.

O'Donovan, G., 2002. Environmental disclosures in the annual report: Extending the applicability and predictive power of legitimacy theory. Account Audit Accountab J 15(3), 344-371.

Orij, R., 2010. Corporate social disclosures in the context of national cultures and stakeholder theory. Account Audit Accountab J 23(7), 868-889.

Orlitzky, M., Schmidt, F.L., Rynes, S.L., 2003. Corporate social and financial performance: A meta-analysis. Organ. Stud. 24(3), 403-441.

Patten, D.M., 1991. Exposure, legitimacy, and social disclosure. J Account Public Pol 10(4), 297-308.

Patten, D.M., 2002. The relation between environmental performance and environmental disclosure: A research note. Account Org Soc 27(8), 763-773.

Perego, P., Kolk, A., 2012. Multinationals' accountability on sustainability: The evolution of third-party assurance of sustainability reports. J Bus Ethics 110(2), 173-190.

Plumlee, M., Brown, D., Marshall, R.S., 2015. Voluntary environmental disclosure quality and firm value: Further evidence. J. Account. Public Policy 34(4), 336-361.

RepRisk, 2016. Our approach. https://www.reprisk.com/our-approach (accessed 13.03.16).

Roberts, R.W., 1992. Determinants of corporate social responsibility disclosure: An application of stakeholder theory. Account Org Soc 17(6), 595-612.

Ruhnke, K., Gabriel, A., 2013. Determinants of voluntary assurance on sustainability reports: An empirical analysis. J Bus Econ 83(9), 1063-1091.

S&P Dow Jones Indices, RobecoSAM, 2014. Dow Jones sustainability indices methodology.

Sethi, S.P., 1978. Advocacy advertising – The American experience. Calif. Manage. Rev. 21(1), 55-67.

Stanny, E., Ely, K., 2008. Corporate environmental disclosures about the effects of climate change. Corp. Soc. Resp. Env. Manage. 15(6), 338-348.

Suchman, M.C., 1995. Managing legitimacy: Strategic and institutional approaches. Acad Manage Rev 20(3), 571-610.

Tukey, J.W., 1962. The future of data analysis. Ann. Math. Stat. 33(1), 1-67.

Ullmann, A.A., 1985. Data in search of a theory: A critical examination of the relationships among social performance, social disclosure, and economic performance of U.S. firms. Acad Manage Rev 10(3), 540-557.

van der Laan Smith, J., Adhikari, A., Tondkar, R.H., 2005. Exploring differences in social disclosures internationally: A stakeholder perspective. J Account Public Pol 24(2), 123-151.

Verrecchia, R.E., 1983. Discretionary disclosure. J Account Econ 5(1), 179-194.

Verrecchia, R.E., 2001. Essays on disclosure. J Account Econ 32, 97-180.

White, H., 1980. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. Econometrica. 48(4), 817-838.

Wiseman, J., 1982. An evaluation of environmental disclosures made in corporate annual reports. Account Org Soc 7(1), 53-63.

Wood, D.J., 1991. Corporate social performance revisited. Acad Manage Rev 16(4), 691-718.